UNCLASSIFIED

Defense Technical Information Center
Compilation Part Notice

ADP014614

TITLE: Precision of Simultaneous Measurement Procedures

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: Proceedings of the Eighth Conference on the Design of Experiments in Army Research Development and Testing

To order the complete compilation report, use: ADA419759

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP014598 thru ADP014630

UNCLASSIFIED

PRECISION OF SIMULTANEOUS MEASUREMENT PROCEDURES

W. A. Thompson, Jr.
National Bureau of Standards
Washington, D. C.

1. ABSTRACT. We consider the problem of measurement under the following conditions: The process of gathering the data is such that on any given item only one opportunity for measurement occurs, but it can be observed simultaneously by several instruments. The items to be measured are variable so that one cannot obtain replicate observations with the same instrument which would show directly the variance of the instrument readings. Procedures are discussed for estimating the precisions of the instruments and the variability of the items being measured.

An example due to Simon and Grubbs is helpful in fixing ideas. The burning times of thirty similar fuzes are determined by several different observers. We limit our discussion to the data taken by observers A and C; hence there are two determinations of the burning times of thirty different fuzes or sixty observations in all. If each of the fuzes had the same running time (which is the manufacturer's goal) and if both of the observers were absolutely accurate, then all sixty observations would be equal. However, considerable inequality in such data always occurs due to variation in the manufacturing process and inaccuracy of the observations. It then becomes desirable to use the sixty observations to answer as many questions as possible about measurement bias and precision, mean fuze running time, and variability of burning times about their mean.

2. THE MODEL. With the verbal description of the previous section in mind, consider the following mathematical formulation. Let $x_1, \ldots, x_N$ denote the true values of the items to be measured. Assume that $x_1, \ldots, x_N$ constitute a random sample of size N selected from a population having mean μ and variance σ^2. Each of the items in this sample is then measured by p instruments. y_{ij} is the measurement of the i^{th} item $(i=1, \ldots, N)$ according to the j^{th} instrument $(j = 1, \ldots, p)$. The consequence of this measurement is that an instrumentation error e_{ij}, chosen at random from the j^{th} instrument's population of errors, is added to the true value of the i^{th} item:

$$y_{ij} = x_i + e_{ij}. \tag{2.1}$$

The instrumentation errors are taken to be uncorrelated among themselves and also uncorrelated with the items selected for measurement. The mean and variance of e_{ij} are β_j and σ_j^2, respectively; β_j may be called the bias of the j^{th} instrument.

Denoting the vector $(y_{i1}, \ldots, y_{ip})$ by Y_i, we may think of $Y_1, \ldots, Y_N$ as constituting a sample of size $N = n+1$ from a p-variate distribution with mean vector $(\mu + \beta_1, \ldots, \mu + \beta_p)$ and dispersion matrix

$$\Sigma = \begin{pmatrix} \sigma^2 + \sigma_1^2 & \sigma^2 & \cdots & \sigma^2 \\ \sigma^2 & \sigma^2 + \sigma_2^2 & \cdots & \sigma^2 \\ \vdots & \vdots & & \vdots \\ \sigma^2 & \sigma^2 & & \sigma^2 + \sigma_p^2 \end{pmatrix} . \tag{2.2}$$

Notice, in passing, that if all instrument variances are equal, then the model becomes a completely random one-way layout and may be analysed by the methods which appear, for example, in Scheffé [5].

A paragraph on notation will perhaps be helpful. ξ_j will be used as a more succinct notation for $\mu + \beta_j$, $j = 1, \ldots, p$. We will frequently write $\Sigma = (\sigma_{jj'})$ when we mean that $\sigma_{jj'}$ is the element in the j^{th} row and j'^{th} column of Σ. In the same spirit $\Sigma^{-1} = (\sigma^{jj'})$, $A = (a_{jj'})$ and $S = (s_{jj'})$ will be common notations. Here

$$a_{jj'} = \sum_{i=1}^{N} (y_{ij} - \bar{y}_{\cdot j})(y_{ij'} - \bar{y}_{\cdot j'}) , \tag{2.3}$$

and $s_{jj'}$ is the usual unbiased estimate of $\sigma_{jj'}$, i.e. $s_{jj'} = a_{jj'} / n$.

3. POINT ESTIMATES. In [2], Grubbs recommends certain estimates of item and instrument variance. For p = 2 instruments, these estimates are

(3.1) $\sigma^2 \sim s_{12}$, $\sigma_1^2 \sim s_{11} - s_{12}$ and $\sigma_2^2 \sim s_{22} - s_{12}$,

where $\sim$ is to be read "is estimated by". For $p \geq 3$, Grubbs recommends

$$\sigma^2 \sim \frac{2}{p(p-1)} \sum_{j<j'} s_{jj'},$$

(3.2) $$\sigma_1^2 \sim s_{11} - \frac{2}{p-1} \sum_{j=2}^{p} s_{1j} + \frac{2}{(p-1)(p-2)} \sum_{2 \leq j<j'} s_{jj'},$$

with an analogeous estimate of the other instrument variances. Gaylor [1] shows that Grubbs' estimate of σ^2 is equivalent to a familiar variance component estimate. These estimates are reasonable in that they have the correct dimensionality, are unbiased, and have appropriate symmetry properties when the instrument labels are interchanged. Further, if the underlying distribution is normal, then Grubbs' estimates are simple functions of the sufficient statistics and in the case $p=2$ they have a maximum likelihood property. However, as Grubbs has verbally pointed out his estimates are unreasonable in that they frequently assume negative values even though the parameters themselves must be non-negative by their very definition.

For $p=2$ this objectional characteristic has been eliminated in [8]; here the altered estimates of table 1 have been proposed. The top line of this table yields the estimates (3.1) under conditions where they are non-negative. The remaining entries show how Grubbs' estimates can be modified when negativity would result from using (3.1). These modified estimates have been derived, under normality assumptions, from a principle of restricted maximum likelihood which is fairly well accepted in other brances of statistical practice. A tilda placed over a parameter indicates its restricted maximum likelihood estimate.

Table 1. Non-negative estimates in the two instrument-case.

Conditions	$\tilde{\sigma}^2$	$\tilde{\sigma}_1^2$	$\tilde{\sigma}_2^2$
$s_{11} \geq s_{12}$ $s_{22} \geq s_{12} \geq 0$	s_{12}	$s_{11} - s_{12}$	$s_{22} - s_{12}$
$s_{22} \geq s_{12} > s_{11}$	s_{11}	0	$s_{11} + s_{22} - 2s_{12}$
$s_{11} \geq s_{12} > s_{22}$	s_{22}	$s_{11} + s_{22} - 2s_{12}$	0
$s_{12} < 0$	0	s_{11}	s_{22}

4. RELATIVE PRECISION. It is clear that if the instrumentation of an experiment is to be effective then the instruments must be precise relative to the variability of the quantity being measured. A frequently quoted rule of thumb is that the instrument precision should be an order of magnitude greater than that of the item being measured. Such a statement has no firm meaning unless a measure of instrument precision and a measure of item variability have been agreed upon. Here we adopt $\Delta_1 = \sigma / \sigma_1$ as a measure of the relative precision of the first instrument. Then, for example, the above mentioned rule of thumb would become $\Delta_1 \geq 10$.

In the two-instrument case, assuming normality, we may use a result of Roy and Bose [3] to make inferential statements of a statistical nature concerning the parameter Δ_1. In our terminology their result states that

$$a_{11} \left[\frac{n-1}{|A|} \right]^{\frac{1}{2}} \left(\frac{a_{12}}{a_{11}} - \frac{\sigma_{12}}{\sigma_{11}} \right) \tag{4.1}$$

has the t-distribution with n - 1 d.f. where $|A| = a_{11}a_{22} - a_{12}^2$. Noting that $\sigma_{12}/\sigma_{11} = (1 + \Delta_1^2)^{-1}$, we may verify that the quantity (4.1) is less than t_α, if and only if

$$\Delta_1^2 > \frac{a_{12} - t_\alpha \left(\frac{|A|}{n-1} \right)^{\frac{1}{2}}}{a_{11} - a_{12} + t_\alpha \left(\frac{|A|}{n-1} \right)^{1/2}} . \tag{4.2}$$

Hence if t_α is the upper α percentage point of the t- distribution with n-1 d.f. then the square root of the right hand side of (4.2) provides a lower confidence bound for Δ_1, the confidence coefficient being $1-\alpha$. The inequality (4.2) can also be used for the purpose of hypothesis testing For example, we may reject $\Delta_1 \geq 10$ at the significance level α if (4.2) is violated with $\Delta_1^2 = 100$.

5. A SIMULTANEOUS CONFIDENCE REGION. For some purposes it may not be enough to consider relative precision; we may be interested in the actual non-relative precisions and the item variability. Estimation of

the parameters $\sigma^2, \sigma_1^2, \ldots, \sigma_p^2$ was discussed in section 3; but how reliable are estimates? This question is dealth with in [9], again under assumptions of normality.

In the two-instrument case, the probability is at least $1 - 2\alpha$ that the following three relations hold simultaneously

$$
(5.1)\quad
\begin{aligned}
&\left|\sigma^2 - a_{12} K\right| \leq M(a_{11}a_{22})^{\frac{1}{2}} , \\
&\left|\sigma_1^2 - (a_{ii} - a_{12})K\right| \leq M\left[a_{ii}(a_{11}+a_{22}-2a_{12})\right]^{\frac{1}{2}}; \quad i = 1, 2.
\end{aligned}
$$

Here K and M are to be found in Table 2 under the desired value of 2α.

Table 2. The table gives values of K and M which yield 1 - 2α confidence regions when used in conjunction with the relations (5.1).

2α	.01		.05	
n	K	M	K	M
3	99.78	99.72	19.79	19.71
4	12.38	12.33	4.146	4.077
5	3.980	3.931	1.726	1.665
6	1.903	1.853	.9636	.9088
7	1.120	1.078	.6290	.5786
8	0.7459	.7076	.4516	.4032
9	0.5389	.5031	.3453	.3022
10	0.4120	.3782	.2761	.2357
11	0.3282	.2963	.2280	.1901
12	0.2698	.2395	.1932	.1573
13	0.2272	.1983	.1663	.1328
14	0.1951	.1675	.1464	.1140
15	0.1702	.1438	.1301	.09925
16	0.1505	.1251	.1169	.08738
17	0.1344	.1100	.1060	.07767
18	0.1213	.09772	.09632	.06962
19	0.1103	.08752	.08904	.06287
20	0.1009	.07896	.08237	.05713
22	.08610	.06546	.07152	.04795
24	.07484	.05538	.06311	.04098
26	.06605	.04763	.05641	.03554
28	.05901	.04152	.05096	.03121
30	.05328	.03660	.04644	.02768
35	.04272	.02778	.03796	.02127
40	.03556	.02200	.03205	.01700
45	.03040	.01797	.02771	.01398
50	.02652	.01503	.02440	.01176
60	.02109	.01110	.01967	.00875
70	.01748	.00862	.01646	.00684
80	.01492	.00694	.01415	.00553
90	.01300	.00575	.01241	.00460
100	.01152	.00486	.01104	.00390

Page Blank

For more than two instruments a similar result is valid. We have the following relations with probability at least $1 - 2\alpha$

(5.2)
$$\max_{j \neq j'} \left[a_{jj'} K - M (a_{jj} a_{j'j'})^{\frac{1}{2}} \right]$$
$$\leq \sigma^2 \leq \min_{j \neq j'} \left[a_{jj'} K + M (a_{jj} a_{j'j'})^{\frac{1}{2}} \right] ,$$
$$\max_{j \neq 1} \left\{ (a_{11} - a_{1j}) K - M \left[a_{11} + a_{jj} - 2a_{1j} \right]^{\frac{1}{2}} \right\}$$
$$\leq \sigma_1^2 \leq \min_{j \neq 1} \left\{ (a_{11} - a_{1j}) K + M \left[a_{11} + a_{jj} - 2a_{1j} \right]^{\frac{1}{2}} \right\},$$

plus $p - 1$ similar inequalities involving $\sigma_2^2, \ldots, \sigma_p^2$. Unfortunately, for p in excess of two, tables of K and M are unavailable. The only result which is currently ready for use is an approximation valid for large n: Choose ℓ to satisfy $P(\ell \leq \chi^2_{n-p+2}) = 1 - 2\alpha$, write $K = M = 1/2\ell$ and substitute this common value in (5.2). I feel obliged to point out that for $p = 2$, the only case where exact values are available, this approximation is rather poor.

6. NUMERICAL EXAMPLE. Returning to the fuze burning time data, we may identify observer A as the first instrument and observer C as the second. From Table I of Grubbs' paper [2] we obtain $a_{11} = 1.3671$, $a_{22} = 1.3227$, $a_{12} = 1.3320$ and $n = 29$. From the third row entry of our table 1, we estimate $\tilde{\sigma} = .21$, $\tilde{\sigma}_1 = .03$ and $\tilde{\sigma}_2 = 0$. By the method of section 4 we obtain, for example, that the relative precision of observer C exceeds 5.1 with a confidence of 95%. Alternatively, from a hypothesis testing point of view we would reject the rule of thumb requirement, $\Delta_2 \geq 10$, at the 5% level. The relations (5.1) and table 2 yield the following 95% simultaneous confidence region: $.16 \leq \sigma \leq .32$, $0 \leq \sigma_1 \leq .09$ and $0 \leq \sigma_2 \leq .07$. In calculating these simultaneous confidence intervals we have replaced all negative lower bounds by zero. Notice that the confidence intervals bracket their respective estimates and hence, in the confidence region sense, indicate the uncertainty of these estimates.

I would like to express my appreciation to the National Academy of Sciences-National Research Council for granting me the opportunity to participate in their postdoctoral research program, and in particular to the Natïonl Bureau of Standards for its support of my research under this program. Also I wish to thank Mrs. Marion Carson for her help in carrying out the computations for table 2.

REFERENCES

[1] D. W. Gaylor, "Equivalence of two estimates of product varinace", J. Amer, Stat, Assoc., vol. 51 (1956), pp. 451-453

[2] Frank E. Grubbs, "On estimating precision of measuring instruments and product variability", J. Amer. Stat. Assoc., vol 43 (1948), pp. 243-264.

[3] S. N. Roy and R. C. Bose, "Simultaneous confidence interval estimation," Ann. Math. Stat., vol. 24 (1953), pp. 513-536.

[4] Thomas S. Russel and Ralph Allan Bradley, "One-way variances in a two-way classification", Biometrika, vol. 45 (1958), pp. 111-129

[5] Henry Scheffe, The Analysis of Variance, John Wiley and Sons, Inc., N. Y., 1959.

[6] L. E. Simon, "On the relation of instrumentation to quality control", Instruments, vol. 19 (1946), pp. 654-656.

[7] H. Fairfield Smith, "Estimating precision of measuring instruments", J. Amer. Stat. Assoc., vol. 45 (1950), pp. 447-451.

[8] W. A. Thompson, Jr., "The problem of negative estimates of variance components", Ann. Math. Stat., vol 33 (1962), pp. 273-289.

[9] W. A. Thompson, Jr., "Estimation of dispersion parameters", in preparation.